道德之一元论与多元论：缘起、内涵与论争*

张昊天[1] 喻 丰[1] 许丽颖[2] 玄哲理[1]

([1]武汉大学哲学学院心理学系, 武汉 430072)

([2]清华大学马克思主义学院, 北京 100084)

摘 要 一元论与多元论之争是道德心理学领域近些年最为活跃的理论交锋之一。道德一元论认为所有外在的道德现象与内在的道德结构都可以用一种因素来解释。道德多元论则认为道德不能只用单一因素来解释，而是存在多个不同质的道德维度，且具有文化敏感性。对应道德理论和道德基础理论是这场争议的典型代表。双方就伤害的解释力、道德失声现象、模块化道德与洁净维度独立性等方面展开论争。未来研究应从三个具体方面进一步探索道德之一元论与多元论难题，进而保持道德心理学领域的理论活力。

关键词 一元论，多元论，道德基础理论，对应道德理论，海特-格雷之争

1 前言

“情感或理性？”“客观或相对？”“先天或后天？”，道德心理学领域存在诸多争议与分庭抗礼的理论。近年来，该领域最为活跃的理论交锋当属道德一元论(moral monism)与多元论(moral pluralism)之争(Graham et al., 2018)。一元论和多元论的分野在早期道德哲学思想中已有体现。亚里士多德创立的美德伦理学主张，人之至善即为“中道(Mesotês)”。中道德性是理智与伦理、相对与绝对的统一，符合中道则为善，过与不及就是恶。中道的最终归宿是实现社会的公平正义(亚里士多德, 2003)。罗尔斯(John Rawls)亦认为道德的核心意蕴即为正义，目的是保证公民的权利和社会合作的公正(Rawls, 1971)。此即一元论思想，将道德视为某种单一维度，以解释所有外在的道德现象和内在的道德本质。当今道德心理学界的一元代表理论有道德发展阶段论(stages of moral development)(Kohlberg, 1969)，对应道德理论(dyadic morality theory)[1] (Schein & Gray, 2018b)等。反观多元论则主张道德具有多个维度，这些维度在不同的社会文化背景下得以体现。社会学先驱涂尔干(Émile Durkheim)认为，道德具有三大基本要素：纪律精神、牺牲精神和自律精神(Durkheim, 1925/1961)。当今道德心理学界的

收稿日期：2021-12-01

*国家社科基金青年项目(20CZX059); 国家自然科学基金青年项目(72101132)

通信作者：喻 丰, E-mail: psychpedia@whu.edu.cn

许丽颖, E-mail: liyingxu@mail.tsinghua.edu.cn

[1] 笔者认为“dyadic morality theory”较为准确贴切的翻译应为“对应道德理论”。在 Merriam-Webster 词典里，dyadic 第一个释意是“pair”，即一种一对一关系。其次，Gray, Young 等(2012)认为道德发生的过程由道德主体与道德客体形成对应关系，缺一则道德判断无法发生。之前有学者将此术语翻译为“道德二元论”，认为此理论是一种二元论视野下的道德判断(詹泽, 吴宝沛, 2019)。然而，这种翻译方式似不妥当，对应道德理论非但不是二元论，反而是彻底的一元论，即所有道德判断的发生都是由感知伤害引起的。

代表理论有三元道德话语理论(triadic moral discourse theory)(Shweder et al., 1997)、关系模式理论(relational model theory)(Rai & Fiske, 2011)、道德基础理论(moral foundations theory)(Graham et al., 2018)等。

一元论与多元论之争是涉及到道德心理学本体论层面的理论探讨，即道德是否可能化约为某种单一维度，还是存在多种不同质的道德价值？这对于认识道德本质有着重要意义。同时，从理论建构的层面来讲，好的理论模型是否应该尽可能多地涵盖道德元素，解释复杂的道德现象，还是应该遵循"奥卡姆剃刀原则（即'如无必要，勿增实体'）"，追求理论的简约性？"一元还是多元"的讨论同样对此问题有着重要的价值。

在以上诸多理论中，以库尔特·格雷(Kurt Gray)为首的对应道德理论和以乔纳森·海特(Jonathan Haidt)为首的道德基础理论是这场争议的典型代表。近 5 年来，双方多次就道德究竟是一元还是多元的问题发表理论或实证文章进行探讨(e.g., Graham, 2015; Gray, Waytz, et al., 2012; Gray & Keeney, 2015a, 2015b; Schein & Gray, 2015)，其理论交锋也延续至今，尚无定论。本文旨在对道德一元论与多元论的争议进行全面系统地总结，并着重考察近些年来十分活跃的对应道德理论与道德基础理论的交锋。未来研究应从明晰道德结构、结合一元论与多元论研究各自的优势建构更为完善的整合式模型以及加强对"非怪异(Non-WEIRD)"文化群体尤其是中国本土文化的关注等方面，进一步推动对于道德本质的探索。

2 道德一元论与道德多元论：内涵与证据

2.1 道德一元论

2.1.1 定义与代表性理论

道德一元论认为所有表面的道德现象与内在的道德结构都可以用一种因素来解释(Beal, 2020)。道德心理学先驱柯尔伯格(Lawrence Kohlberg)认为道德可以归结为一种因素，即关乎正义的推理，并且认为道德发展有一个明确的终点，即成为一个主张公平与平等的道德思想家(Kohlberg, 1969)。除此之外，格雷等人认为道德内涵的多元只是一种外在表现形式, 其核心因素只有一个，即感知伤害(Gray, Waytz, et al., 2012)。

在以上关于道德一元论的理论中，近些年最活跃、最有代表性的当属格雷等人提出的对应道德理论。该理论认为道德态度和信念的多元化可以完全用感知伤害(perceived harm)来解释(Schein & Gray, 2018b)。在道德判断的过程中，存在一个"认知范本(cognitive template)"，即不道德行为是由有意图的道德主体(intentional moral agent)向遭受痛苦的道德客体

(vulnerable moral patient)施加伤害的过程，且两者之间存在因果关系(causation)，这些是感知伤害必不可少的三种因素(Schein & Gray, 2018b)。道德主体意指通过自身的认知与推理形成主观判断，并可以根据自身意图实施道德行为且能够承担责任的道德行动者，例如通常社会意义中的“施害者(perpetrator)”；而道德客体则是行动者所实施道德行为的被动接受者，例如通常社会意义中的“受害者(victim)”(Gray & Wegner, 2009)。重要的是，产生感知伤害与不道德判断的前提在于有意图做出施害行为的道德主体与受害的道德客体之间须具备因果关系，且两者的因果关系越清晰，伤害就越明显；当因果关系不明确时，例如道德两难困境中常出现的“副作用效应(side-effect effect，意指某种行为带来的附加影响)”，道德判断的严苛性就会减弱(Schein & Gray, 2018b)。

此外，“感知伤害”由于其丰富的内涵，需要进一步予以澄清。首先，格雷等人认为“伤害”并非客观、外在的，例如有些被人们认为不道德的事件（如与冻鸡交媾、有保护措施的兄妹乱伦等），虽然没有造成直接的、客观的身体伤害或财产损失，但仍被人们知觉为存在对身体与社会规范的潜在伤害(e.g., Royzman et al., 2015; Schein & Gray, 2016)。其次，研究表明人们在极短时间内就能知觉到伤害并予以归类，说明对伤害的感知是直觉的，而非理性的(Schein & Gray, 2018b)。最后，伤害的表现形式是多种多样的，不仅表现为常见的身体或情感伤害，同时包括欺骗伤害、群体伤害、不敬伤害、灵魂伤害等(Hartman et al., 2022)。虽然伤害在表现形式上支持“多元论”，但从其本质而言，对应道德理论认为所有道德判断都可以归结为一种因素——感知伤害(Schein & Gray, 2018a)。

感知伤害又是通过何种认知机制得以实现？Gray, Young 等人(2012)认为这依据于对他人的心智感知(mind perception，即对人与物等实体是否存在心智能力的感知)，而这亦是道德的本质。通过“自下而上（bottom-up approach）”的数据驱动方法，Gray 等人(2007)利用因素分析法将人们对心智的感知维度分为两部分：能动性(agency)与感受性(experience)。能动性关乎行动与思考，对应的是道德责任；而感受性则关乎知觉与情绪，对应的是道德权利。道德主体通常被知觉为能动性大于感受性，而道德客体则相反(Gray & Wegner, 2009)。正如违反法律的精神病患者往往无须承担刑事责任，原因之一就是精神病患者缺乏心智能力，从而无法成为道德主体。需要说明的是，这两个维度并非道德维度，而是个体对于心智的感知，用于解释道德判断中的“认知范本”。对应道德理论主张的道德一元论维度只有一个——感知伤害。

2.1.2 支持证据

伤害在道德判断中的重要作用毋庸置疑，格雷等人试图证明感知伤害为道德判断的最主

要影响因素，甚至是唯一的影响因素，而非以往道德心理学中所认为的其他因素。以厌恶(disgust)为例，格雷认为其本身很难成为道德化的成因，因为许多行为唤起厌恶并非是不道德的，例如苍蝇飞进米饭中(Gray & Schein, 2016)。在 Schein, Ritter 等人(2016)的研究中，研究者请被试评价 24 种厌恶相关行为的道德程度、伤害程度和厌恶程度，这些行为包括令人厌恶的性行为（如口交）、和病原体有关的厌恶行为（如看到蟑螂在地板上跑窜）以及道德违背行为（如欺骗朋友）。结果表明，感知伤害可以最好地区分令人厌恶的不道德行为与仅仅是厌恶的行为。在另外两项研究中，研究者请被试评价同性恋婚姻和亵渎神灵的不道德程度，厌恶感虽然能直接预测不道德程度，但感知伤害在其中起完全中介作用，也就是说，感知伤害才是解释不道德行为的最强有力因素，当其存在时，厌恶感与不道德判断之间没有直接的因果关系(Schein, Ritter, et al., 2016)。

除了直接从“感知伤害”本身进行论证，格雷等人还通过一系列实验，证明“洁净(purity)”道德本身与感知伤害并无明显差别。研究者请被试对三种行为的不道德和危害性进行评分，包括伤害违背行为（如谋杀）、无害行为（如乘坐公交车）以及洁净违背行为（如在圣经上涂粪便）。结果表明，当人们被迫迅速做出反应时，对伤害的感知会更高，即使在一些无客观伤害的洁净道德情境下，被试仍然直觉地认为这些行为存在不同程度的伤害。这些结果表明，感知不道德的伤害是通过直觉，而无须刻意地合理化(Gray et al., 2014)。

此外，在政治领域中，自由主义者(liberals)和保守主义者(conservatives)有着不同的道德心理基础。例如 Graham 等人(2009)发现自由主义者更看重“关爱”、“公平”等个体化道德基础，而保守主义者更看重“忠诚”、“权威”、“洁净”等绑系性基础。但是格雷等人认为，不同政治倾向之间并没有不可逾越的道德鸿沟。相反，他们的研究表明，无论是自由派抑或是保守派，在本质上有着相同的道德观念，那就是对于伤害的关注，区别只在于看待伤害的方式不同而已(Landy, 2016; Schein & Gray, 2015)。通过 7 个研究，研究者发现无论是倾向于保守派还是自由派的被试，在回忆起不道德事件时，最先想到的都是关于伤害的事件（实验 1），而且都认为伤害是不道德事件的最重要因素（实验 2、实验 3）。此外，通过对内隐测试指标（反应时）的测量，发现伤害与不道德事件紧密相联（实验 4）。研究者还请被试评价道德基础理论中常用的 10 个道德事件，发现不同的道德基础之间存在高度相关（实验 5），且伤害对不道德事件的预测最强（实验 6），甚至一些之前被认为是洁净违背的事件（例如伴侣出轨）也与伤害有更加紧密的联系（实验 7）(Schein & Gray, 2015)。

根据以上证据，格雷等人认为伤害才是解释道德判断最有力的因素，其他学者所认为的洁净、公平、忠诚等道德基础，均可以通过感知伤害来解释(Schein & Gray, 2015)，而不同政

治派别的道德基础偏好，本质上都是针对伤害的不同解读。

2.2 道德多元论

2.2.1 定义与代表性理论

道德多元论认为道德不能只通过单一因素来解释，而是存在诸多不同质的维度，这些维度是在不同文化和社会阶层的人群中呈现的，具备文化敏感性(cultural sensitivity)(Graham et al., 2013, 2018)。针对柯尔伯格的一元论道德观，早期的多元论支持者卡罗尔·吉利根(Carol Gilligan)与埃利奥特·图列尔(Elliot Turiel)认为道德的核心不仅在于追求正义，亦在于关怀他人(Gilligan, 1988)与避免伤害(Turiel, 1983)。文化人类学家理查德·施韦德(Richard Shweder)是最早较为系统地提出道德多元论主张的学者。基于在印度的实地调查，施韦德和同事批评前人的一元论道德观只是反映了西方世俗文化(secular culture)下广泛持有的个体主义道德观，忽视了其他文化可能存在的诸多道德领域。基于全球性的文化背景，施韦德提出更为多元的“三元道德话语理论”，该理论认为观察到的道德领域主要可以分为三种：其一为“自治伦理(the ethic of autonomy)”，即人们依靠诸如伤害、权利和正义等观念来保护个体的自主性，对应的道德情绪为愤怒；其二为“社群伦理(the ethic of community)”，即人们依靠责任、尊重和忠诚等观念来维护制度和社会秩序，对应的道德情绪为轻蔑；其三为“神性伦理(the ethic of divinity)”，即人们依靠诸如圣洁和罪恶等观念来保护每个人内在的神性，维护个人洁净与圣洁的道德，对应的道德情绪为厌恶(Shweder et al., 1997; Rozin et al., 1999)。

此外，文化人类学家阿兰·费斯克(Alan Fiske)对一元论道德也持批判态度。基于早年创建的“关系模式理论(relational models theory)”，费斯克认为道德是一种“关系规范(relationship regulation)”，以维护与他人的社会合作关系，其中有四种不同的基础道德动机，即“团结(unity)”、“阶层(hierarchy)”、 “平等 (equality)”和“比例(proportionality)”：团结是通过维护内群体完整性来避免或消除消极事件的威胁；等级制度是社会群体中尊重上级的动机；平等指维护社群平等的福利与发言权；比例指奖励和惩罚须与功绩成比例，利益须与贡献成比例(Rai & Fiske, 2011)。

除了上述道德理论，近些年道德多元论的代表性理论当属道德基础理论(Graham et al., 2013, 2018)，旨在解释道德判断的多样性和相对性。道德基础理论主张先天主义(nativism)、文化学习(cultural learning)、直觉主义(intuitionism)和多元主义(pluralism)。先天主义认为人类已经在进化过程中形成了先天的道德基础，经由后天经验不断修正以形成成熟的道德价值观；文化学习意指先天道德基础在文化进化过程中不断发展；直觉主义意指道德判断基于直

觉而非沉思。最后，道德基础理论主张多元主义，因人类在进化史上面临诸多适应性社会挑战，因此在应对挑战时形成了不同的道德基础。

迄今为止，大多数研究都集中在五种道德基础上，前两种是个体化基础(individualizing foundations)，强调对个体权利的保护。其一为*“关爱/伤害(Care/harm)”*：在进化过程中，那些对孩子需求和痛苦更敏感的看护者，其后代更有可能生存下去。此道德基础意味着对自己与他人的同情和关怀，以及对痛苦和伤害的避免。其二为*“公平/欺骗(Fairness/cheating)”*：所有群居动物都有其社会性，为了避免两败俱伤的局面，需要公平与互惠利他来实现共同利益，并且需要监督他人的行为和声誉，形成惩罚欺骗者的社会规范。

后三种是绑系性基础(binding foundations)，强调的是约束和建立社会群体的价值观。其一为*“忠诚/背叛(Loyalty/betrayal)”*：在抢夺社会资源的过程中，最具凝聚力的联盟往往会战胜凝聚力较弱的竞争联盟，所以忠于内群体（例如爱国）也是社会所规定的美德。其二为*“权威/颠覆(Authority/subversion)”*：拥有权力的族群会受到尊重并被视为合法，通常比无领导或无规范的族群运作得更好。不尊重权威或传统的人往往会引发人们对无序的焦虑和对破坏者的蔑视。其三为*“洁净/堕落(Purity/degradation)”*：是人类为了远离疾病感染而进化出的道德基础，并伴随着对病原体特有的厌恶情绪。当今时代的洁净基础也泛化到社会生活中，包括拒绝外群体和排斥不遵守群体神圣习俗的人。

尽管大多数研究都集中在这五个道德基础上，但可能还有许多其他的道德基础；例如，自由/压迫(liberty/oppression)、诚实/说谎(honesty/lying)等(Graham et al., 2018)。不过，道德基础理论虽然经过多次修订，目前最为公认的仍然是以上提到的五种道德基础。

2.2.2 支持证据

道德多元论均有丰富的实证研究为依据。基于印度的田野调查，施韦德等人首先通过访谈研究确定基本的道德元素（例如，男尊女卑），再采用层序聚类分析的方式从 25 个基本元素中提取出三个因子，分别为自主伦理、神性伦理与社群伦理 (Menon & Shweder, 1994)。而 Rozin 等人(1999)在此基础上提出 CAD(contempt, anger, disgust)模型，旨在寻找三种伦理所对应的道德情绪，并请被试将三种伦理对应的道德情境与情绪图片或词汇进行匹配，并做出相应的情绪表情。结果表明自主伦理更多对应愤怒，社群伦理更多对应轻蔑，神性伦理更多对应厌恶，再一次从情绪视角验证了道德价值观的多元性。

此外，同为文化人类学家的阿兰·费斯克则从非洲传统部落文化下进行田野调查，从社会关系的视角研究人的合作与计划等社会行为，并认为存在四种基本关系类型，分别是：公共共享(communal sharing)、权力排序(authority ranking)、平等匹配(equality matching)和市场

定价(market pricing) (Fiske, 1992)，此类型亦通过验证性因素分析得以证实(Haslam & Fiske, 1999)，并在中国、韩国等四个国家验证了跨文化普适性(Fiske, 1993)。Rai 和 Fiske(2011)在此基础上进一步提出关系规范模型，认为道德的目的应是维护有效的社会关系。实证研究发现，无论是在消费者行为(McGraw et al., 2012)、组织行为(Giessner & Van Quaquebeke, 2010)抑或司法(Copes et al., 2013)领域，均能发现人们依赖这四种不同的动机做出道德判断，证明了道德结构与成因的多元化。

而海特等人则是基于道德话语理论、关系模式理论等前人研究建构出道德基础理论，并汲取文化人类学、进化心理学、社会学等学科的智慧，采取的是自上而下的理论建构方法(top-down approach)，在初期主要以理论思辨为主(Haidt, 2012)，并无实证研究探索为何道德可以具体化为五个维度。之后研究者方才通过问卷编制、质性访谈等方法验证道德基础的成立依据。Graham 等(2011)编制了道德基础问卷(Moral Foundations Questionnaire)，施测结果表明该问卷具有良好的内部效度、外部效度与实用效度。验证性因素分析表明，道德五因素结构模型拟合度良好，这为道德基础的成立提供了实证依据(Graham et al., 2011)。跨文化研究结果表明，东方群体对于忠诚和洁净道德基础的关注高于西方群体，而对另外三项道德基础的关注并无明显差异，而且政治倾向与道德基础的关系并未因为文化的不同而产生变化，这为道德多元论的跨文化普适性提供了实证依据(Graham et al., 2011)。后续研究也证明了道德基础问卷具有良好的跨文化稳定性(e.g., Doğruyol et al., 2019; Yalçındağ et al., 2019)。以上研究均说明，在不同的特征的群体中会呈现出不同的道德基础，且这种道德多样性具备跨文化的稳定性，从而支持了道德多元论主张。

另外，海特等人还发现不同政治倾向的人群侧重不同的道德基础。McAdams 等人(2008)通过质性研究，对 128 名深度参与宗教和政治活动的成年人进行了生活叙事采访，并对他们的回答进行了与五大道德基础相关的编码。结果表明保守主义者更为强调尊重社会等级、对内群体的忠诚以及洁净或神圣道德基础，而自由主义者则更重视关于伤害和公平的道德基础。Graham 等人(2009)则使用了四种不同的方法，发现自由派比保守派更看重个体化道德基础，而保守派关注所有五个道德基础，且比自由派更看重捆绑性道德基础。最近的元分析研究表明保守派和自由派的基本差异得到了支持(Kivikangas et al., 2021)，这说明不但存在多种道德基础，其分布也因阶层、政治倾向而异，再次证明了道德基础的多元性。

3 道德一元论与多元论之争

在道德心理学发展早期，无论是以柯尔伯格为代表的道德一元论，还是以施韦德、费斯克等人为代表的道德多元论，只是呈现各自为营的局面，尚未形成理论对话。直至后来道德基础理论和对应道德理论的提出，一元论与多元论才真正形成理论交锋的局面。2012 年，格雷等人在著名理论评论类期刊 *Psychological Inquiry* 发表靶子文章(target article)：《心智感知是道德的本质》(Gray, Young et al., 2012)，正式提出道德一元论的主张，即所有的外在道德现象与内在观念均可以由感知伤害来解释。海特与同事在同期发表评论文章，在肯定对应道德理论所做的理论贡献的同时，对道德一元论的主张进行了驳斥(Graham & Iyer, 2012; Koleva & Haidt, 2012)。此后几年，双方一直就道德究竟是一元还是多元的问题进行了大量深刻的讨论。下文将从道德失声现象、模块化与建构主义、洁净道德维度的独立性等方面总结双方论争的主要观点并做出评述。

3.1 一元论对多元论的驳斥

3.1.1 “道德失声”现象并非无害

尽管有些道德情境与客观伤害无关，例如与冻鸡交媾、拿国旗擦马桶等，但人们仍然直觉地认为这些行为是不道德的，当被询问具体原因时，人们却无充分理由支持该判断或决定，此即“道德失声(moral dumbfounding)”现象(Haidt et al., 2000)，也是海特捍卫其直觉主义道德观的主要论据。其中的核心原因是道德判断过程中存在“洁净违背(purity violations)”或“忠诚违背(loyalty violations)”机制(Scott et al., 2016)，与伤害无关，而这些道德基础是促使道德一元论向多元论发展的重要概念(Gray et al., 2021)。但格雷等人否认这些行为并非完全无害，认为海特将感知伤害狭隘地等同于客观伤害(Guglielmo, 2018)，所谓独立的“洁净、忠诚”等机制，均可以通过感知伤害来解释。另外，格雷认为海特等人选取的道德情境太过“怪异”，在一定程度上限制了道德失声现象的解释力。

对于道德失声情境的“无害性”，已有若干实证研究表明道德失声现象涉及到的情境并非无害。Royzman 和同事(2015)重复了传统的道德失声范式。与以往不同，研究者测量了被试是否真的相信这些行为是无害的，例如问：“你认为朱莉和马克(他们是姐弟)发生性关系而不会损害他们之间的关系的可信度有多高？不会有不良后果的可信度有多高？”结果表明，被试坚持认为乱伦会产生不良后果，且对现实世界有负面影响。在随后的访谈中，即使研究人员不断解释乱伦是无害的，被试仍然持怀疑态度，且有 57.1%的被试认为乱伦给朱莉和马克造成了情感伤害。另一项研究也表明，伤害完全可以预测被试对道德失声现象的不道德判断，哪怕仅仅是未发生的“潜在伤害(potential harm)”(Stanley et al., 2019)。基于以上论据，格雷认为所谓“无害错误(harmless wrongs)”，其实都存在不同程度的伤害。

此外，对于道德失声情境的“怪异性”，格雷等人亦批判道德失声现象涉及到的情境太过于“怪异”，诸如与冻鸡交媾、吃死狗、出卖灵魂、在尿里打滚等行为很明显与日常生活中发生的道德难题相去甚远，这些怪异的情境并不能很好地揭示道德认知的本质，而道德的发展是为了指导我们日常的人际交往和促进合作，海特等人的研究显然背离了这一点(Gray & Keeney, 2015a)。在一项研究(Gray & Keeney, 2015b) 中，被试评价道德基础理论中常用的道德情境，并从严重、怪异、错误、伤害以及不洁五个方面进行评估。被试认为相比于伤害情境，不洁情境更加怪异，严重程度也更低。有趣的是，在伤害情境下，被试不仅对伤害性的评分更高，甚至对不洁的评分也高于不洁情境。接着，研究人员又请被试列举自己认为的不洁与伤害情境，发现被试提名的自然情境更贴近现实，如故意杀人，婚姻不忠等。对自然情境和道德基础情境进行对比，发现被试认为道德基础情境更怪异，且伤害性与严重性远不如自然情境。综上，自然情境能比道德基础情境更好地预测伤害和洁净道德，且伤害对于不洁情境的解释力反而比洁净更高。

3.1.2 对道德模块化的驳斥

道德一元论者对多元论者的驳斥还在于模块化假说这一点。格雷认为道德基础理论明确支持模块化(modularity)，即不同的心理特征对应不同的认知模块，具有区分性(distinctness)和特异性(specificity) (Schein, Hester, et al., 2016)。道德基础模块是存在于动物大脑里的“开关”，当一个特定刺激（例如，某种道德违背）激活了一个特定模块（例如，洁净道德基础），从而会导致某种行为或体验（例如，与洁净相关的道德判断）(Haidt, 2012, p. 123)，并激发出模块对应的特定情绪。

但格雷认为道德基础理论的模块化解释缺乏说服力，因为某种行为可能激活多个模块(例如，婚姻不忠可能同时激活忠诚和伤害的道德基础)，而且不同的道德基础之间存在类似的高度重叠。Schein 和 Gray(2015)请被试从道德基础来评定 10 个道德情境，发现所有道德基础的内部一致性系数高达 0.89。此外，道德基础量表的数据也表明不同的道德基础之间存在很高的相关度，例如权威和忠诚之间的相关性为 0.88，伤害和公平之间的相关性为 0.72，洁净和权威之间的相关性为 0.80(Graham et al., 2011)。这些高相关性结果本身不能说明道德基础理论的分维度具备很高的区分度，特别是不同道德基础间相关性高于基础内的一致性(例如，忠诚与权威的相关性大于与其自身的一致性系数)。

此外，格雷也反对海特等人道德情绪模块化(Haidt, 2012; Rozin et al., 1999)的说法，认为某种道德基础不是激发特定的情绪模块，而是同时引发多种情绪体验：例如伤害违背可能同时激发愤怒和厌恶等情绪(Gray et al., 2017)，也有研究者发现洁净违背则引起更多的微笑反

应和愤怒表情，而不是海特、施韦德等人所认为的厌恶情绪(Franchin et al., 2019)。对此，格雷等人则秉持"建构主义道德情绪观(constructivism of moral emotions) (Cameron et al., 2015)"，认为某种道德情绪并非通过不同道德模块才得以激发，而是由规范违背、消极情感和感知伤害三种基本成分的组合而产生，具有领域普适性(domain-general)。某些道德行为因为触发了以上三种因素，进而产生不同的道德情绪，例如愤怒、厌恶等(Gray et al., 2017)。

3.1.3 洁净不足以成为独立的道德基础

关于洁净是否可以作为独立的道德维度是一元论者与多元论者目前争议最多的问题(Gray et al., 2021)。Gray 等人(2021)系统回顾了 158 篇关于洁净道德的研究，发现几乎没有证据支持洁净可以作为一个独立的心理构念，也不能作为一种独特的道德判断机制。洁净道德本身的定义过于复杂，且至少 9 种存在关于洁净的不同定义，涉及到病原体回避、维持自然秩序等。文献计量统计表明洁净的定义和实际测量存在相当大的异质性。在 135 篇测量了洁净道德的文章中，仅有 19 篇(14%)做到了定义与测量相一致，而两者的分离则不利于理论的发展(Gray, 2017; Gray et al., 2021)。相反，伤害的概念则更为简洁，即对身体或心理造成的损伤(Haslam, 2016)，而且实际测量也往往与定义相一致。

此外，在多个实证研究中，洁净与伤害道德基础的无明显的区分度(Gray et al., 2014; Gray et al., in press; Gray & Keeney, 2015b; Schein & Gray, 2015)。无论是针对道德基础理论情境的道德判断，还是实验被试列举的情境，洁净判断与伤害判断存在显著高度相关：rs ≥ .87(Gray & Keeney, 2015b, study 1)。甚至一些之前研究认定为不洁的情境也能够使个体体验到直觉上的感知伤害(Gray et al., 2014; Gray & Keeney, 2015b)。另有研究表明，不洁净情境更有可能被感知为伤害，而伤害情境则很少会被感知为不洁净(Chakroff et al., 2017)，而且伤害相比洁净与道德判断的联系更紧密(Rosenfeld & Tomiyama, 2021)。

最后，Gray 等人(2021)认为道德心理学家所定义的洁净行为可能也是一种关于伤害的抽象道德指令，这些伤害可能不涉及直接的身体伤害，却是一种更高层次的伤害。具体来讲，"殴打他人"属于一阶伤害(first-order harms)的不道德，而要获得更抽象的高阶道德原则，比如某些饮食习惯或宗教仪式（例如，不吃猪肉），则与一个群体的文化预设有关。关于儿童社会化的研究表明，儿童一开始可能会采用自主伦理（例如偷玩具会伤害其他孩子）来解释偷窃行为，但经过文化宗教社会化的过程可能会让孩子后来用神性伦理（偷窃玷污了上帝的意志）来解释不道德行为(DiBianca Fasoli, 2018)。洁净规范在儿童时期较晚出现，且不像一阶伤害那样普遍，而且这种社会化过程是建立在伤害基础上的。

3.2 多元论对一元论的驳斥

3.2.1 一元论过度简化了道德内涵

Haidt 等人(2015)认为对应道德理论对道德现象与道德理念进行了过度简化，而复杂的道德现象本质上不可仅仅化约为伤害，并暗讽格雷将“奥卡姆剃刀(Occam’s Razor)”运用成了“奥卡姆电锯(Occam’s Chainsaw)” (Koleva & Haidt, 2012)，删减掉了一些非常重要的道德维度，例如洁净、公平等。海特还认为，虽然格雷强调道德一元论，但在对于伤害的定义上，又在坚持“伤害多元论”，这些不同类型的伤害，例如身体伤害、人际伤害，心理伤害等，这无非就是其他的道德维度的不同话语表达而已，所以道德一元论研究已经不可避免地走向了道德多元论(Haidt et al., 2015)。

其次，海特认为一元论者过于强调和夸大伤害在道德判断中的作用，使其理论变得不可被证伪(Koleva & Haidt, 2012)。例如人们在面对一些具有明显伤害的事件（杀人、殴打等）中做出了不道德判断，则可以认为伤害因素是人们做出道德判断的最主要依据，但如果是一些无害的情境（例如采取了安全措施的兄妹性关系），一元论者又将其解释为存在“感知伤害”或“潜在伤害”，这就使“伤害”这个概念的定义过于主观和宽泛，对一切不道德现象的解释均能自圆其说，从而不可被证伪，不足以成为严谨的科学命题(Haidt et al., 2015)。

此外，Haidt 等人(2015)认为一元论研究存在诸多方法学问题。关于政治倾向与道德基础的关系，Schein 和 Gray(2015)虽声称自由主义者和保守主义者在道德认知上没有差异。但其中关于政治倾向的测量却有很大的问题：研究者虽采用李克特 7 点等距量表（1-非常自由主义—7-非常保守主义）来测量人们的政治倾向，但却采用了简单地二分法，错误地将中立政治倾向的个体（即在量表中选择了 4）划分为了保守派，导致不可靠的结果。所以未能发现两组人之间的道德基础差异并没有足够的说服力。另外，Schein 和 Gray(2015)在研究中也仅仅考察了伤害相关词汇和道德判断的联结程度，并没有测量其他的道德维度，例如公平、忠诚等，所以不能说明伤害是解释道德判断的唯一因素。

3.2.2 对于道德模块化的辩护

针对一元论者对于道德基础理论模块化假说的驳斥，Haidt 等人(2015)认为这更多是对道德基础理论的误解。道德基础理论并非是五个“福多式模块(Fodorian modularity，即负责特定心理过程的独立功能模块) (Fodor, 1983)”，而是采用了更为灵活和重叠的“群集模块(massive modularity)”概念，具备“功能专门化(functional specialization)”，这些模块处理特定的刺激和信息输入，以此来适应环境。这些模块并不是完全密闭的，而是不断发展的：人类先拥有道德心灵的“初稿”，然后通过特定的文化学习不断发展。

Graham 等人(2018)进一步解释道，道德基础理论从来没有认为道德判断是由五个截然

不同的模块组成。道德的发展始于一些先天的知识，即“原始诱因(original triggers，例如关爱、公平等)”，这些知识存在于学习模块(learning modularity)中，使得学习特定事物很容易，学习其他事物就很难。而在后天的文化学习中，个体在“即时诱因(current triggers，例如某种特定文化)”的影响下产生许多新的模块，从而形成不同的道德基础。这些模块又与特定的情绪和动机反应联系起来。所以复杂的道德结构是建构与不断发展的，不能用完全独立的 5 个模块来解释，更不能只用一个模块（如感知伤害）来解释。

3.2.3 洁净可以成为独立道德维度的证据

面对道德一元论者的质疑，海特认为洁净作为一个独立的道德维度，有着广阔的跨学科证据和深厚的理论内涵，并结合了宗教学，文化人类学，进化生物学，进化心理学等多学科的智慧。早在多年以前，Haidt 等(1993)通过自编的非伤害情境（拿国旗擦马桶，与死鸡交媾等），发现社会阶层较高的大学生认为这些故事是社会习俗或个人偏好的问题。大多数低社会阶层的学生，对这些行为采取更严苛的道德判断，而且情感反应比伤害评估更能预测道德判断。一些关于具身化道德(embodied morality)的实验研究同样发现不洁造成的生理厌恶会增加道德判断的严苛性(Wheatley & Haidt, 2005)。基于此，海特认为不洁厌恶可以解释人们对于非伤害情境的道德责备。此后，关于洁净道德的实证研究不断增长。Graham 等(2018)在最近的理论文章中总结了洁净可以作为独立道德基础的依据，并阐述其与伤害的本质区别。

首先，伤害和洁净道德在认知层面上存在本质不同。例如频率和重要性对伤害判断的影响程度更大，而对洁净判断的影响因素更小(Rottman & Young, 2019)。洁净和伤害所涉及到的思维风格亦有不用，激活或训练分析式思维可以提升个体对于洁净道德基础的关注，而激活或训练整体性思维提升的则是对于伤害道德基础的关注(Alper & Yilmaz, 2020)。当自己是受害者的时候认为洁净违背在道德上更加错误，当他人是受害者的时候认为伤害违背在道德上更加错误(Dungan et al., 2017)。与伤害判断相比，不洁净判断较少地谴责行为本身，而更多地谴责行为主体(Uhlmann & Zhu, 2014)。此外，意图(intention)对洁净的影响要小于伤害判断(Young & Saxe, 2011)，偶然或有意伤害会在右侧颞顶叶交界处产生不同的激活，而偶然或有意侵犯洁净没有这种区别(Chakroff et al., 2017)，这些不同类型的判断也与不同的面部微表情和神经系统有关(Wasserman et al., 2017)。

其次，洁净道德基础可以预测伤害所不能预测的重要心理现象和行为。例如，洁净可以唯一预测诸如同性恋婚姻、安乐死等问题(Koleva et al., 2012)，而其他道德基础和政治意识形态则不能。洁净还预测了对疫苗接种的争论(Amin et al., 2017)、亲环境态度(Rottman et al., 2015)和现实社会网络中的社交距离(Dehghani et al., 2016)。对不洁净的道德关注比对伤害的

判断更能预测自杀的道德判断(Rottman et al., 2014)。洁净判断比与任何其他道德基础相关的判断更能预测厌恶敏感性(Wagemans et al., 2018)，即使在伤害情境下。此外，关于道德化过程的研究表明，通过“道德震慑(Moral Shock, 例如通过流产胎儿的图片增加对堕胎的道德笃念)”的方式，研究者发现道德化的过程是由厌恶作为中介变量，而不是愤怒或伤害，从而驳斥了对应道德理论(Wisneski & Skitka, 2017)。

4 总结与展望

一元论与多元论的分野在早期道德哲学的体系中已有所体现，可以追溯到“道德相对主义(moral relativism，认为不存在普适的道德原则，任何道德价值都只适用于特定文化边界与个人价值系统之内)”与“道德普遍主义(moral universalism，一种一元伦理学立场，认为存在对所有人都适用的普遍伦理)”的分歧，但近些年这类理论交锋进入了道德心理学的视野(Beal, 2020)。表 1 简要汇总了以往道德一元论与多元论研究的共识与争议点。总体来讲，道德基础理论致力于描绘不同文化下人类道德的全貌。相比之下，对应道德理论则更为简约，是“奥卡姆剃刀原则”的典范，具有理论美感。但双方的理念并非截然不同。正如海特认同“关爱/伤害”是道德基础理论中最为核心的维度(Haidt et al., 2015)，格雷也认同道德基础理论的四种之中的三种基本主张（先天主义，文化学习，直觉主义）(Gray & Keeney, 2015a)，但二者最主要的分歧在于“伤害”是否能够解释所有外在的道德现象以及内在的道德态度、认知与情感过程。我们认为未来关于道德一元论与多元论的研究可以从以下三个具体方面展开。

表 1 道德一元论与多元论的共识与争议总结

	道德一元论	道德多元论
定义	所有表面上的道德现象与内在的道德结构都可以用一种因素来解释。	道德存在诸多不同的维度，这些维度是在不同文化和社会阶层的人群中呈现的，具备文化敏感性。
代表人物	亚里士多德、罗尔斯、柯尔伯格、库尔特·格雷等。	涂尔干、吉利根、图列尔、施韦德、阿兰·费斯克、乔纳森·海特等。
支持依据	① 伤害是解释道德判断最强有力的因素。 ② 伤害更加直觉。 ③ 不同政治倾向的人皆认为伤害是道德认知的核心。	① 道德基础量表有良好的构念效度、区分效度、实用效度等。 ② 不同政治倾向、文化、社会阶层的个体体现出不同的道德基础。
论争要点	① 道德失声现象源于感知伤害和情境怪异性，而不是洁净违背。 ② 几乎没有证据支持模块化道德假说。 ③ 洁净不可作为独立的道德维度。	① 一元论过度简化了道德内涵。 ② 道德基础并非是五个“福多式模块”，而是更为灵活和重叠的“群集模块”。 ③ 有大量证据支持洁净是独立的道德基础。
共识	① “关爱/伤害”是道德的核心维度。 ② 道德基于先天主义，文化学习，直觉主义。	

首先，道德需要一个更加清晰的界定，而这一点或许连伦理学也无法做到(MacIntyre, 1957)。道德究竟是一元还是多元，涉及到道德本质(essence)的具体化，甚至不少道德哲学家认为道德可以没有本质(Taylor, 1978)。如果道德有本质，却无法辨明，则此题无解。而如果道德没有本质，则它只可能在语言使用或者认识上存在意义，那么心理学的操作性定义或者被试给出的平均看法就是道德本质。这就导致“专家洞见(expert insights)”与“民众理论(folk theory)”的鸿沟：研究者与受试者对概念的理解可能不同，例如研究者认为的“无害”并不代表被试同样认为无害(De Villiers-Botha, 2020; Royzman et al., 2015)。更棘手的是，在实际研究中，研究者采用的理论定义与操作定义往往不一致。例如上文中提到的洁净道德，仅有 14%的研究做到了理论定义与操作性定义的统一(Gray et al., 2021)。所以，为保证道德研究的有效性，未来研究应尽可能做到将道德概念定义地更加清晰，应当确保每次只检验一种构念。因为道德现象纷繁复杂，有些研究中常用的道德情境并非只对应特定的道德基础，例如“当面训斥父母”这个道德情境(Graham et al., 2011)同时涉及了“关爱/伤害”与“权威/颠覆”两种不同的维度，通奸行为也同时涉及到伤害、洁净与忠诚道德。如果不保证情境清晰地反应某种道德维度，则研究者很难精确指出哪种道德维度影响了人们的道德判断。

其次，除了关注一元论与多元论本身的争议，我们也需从纷争中抽离出来，客观看待这两种道德取向研究的不同特点，彼此取长补短，以推动道德心理整体研究的发展。具体而言，一元论重视道德构念的简洁度以及测量的精确性，多元论则着重强调不同文化下的人群对于道德本质的理解。这是两种不同的理论建构和看待道德本质的方式。但一元论无法解释所有的道德情境，例如因不作为带来的伤害和涉及伤害的公平竞争情境（如拳击比赛）就难以通过对应道德理论来解释(Piazza et al., 2019; Royzman & Borislow, 2022)。同样，多元论研究也很难在不同阶层、不同文化群体中重复出稳定一致的五种道德基础(e.g., Curry et al., 2019; Harper & Rhodes, 2021)。未来研究应该结合一元论与多元论研究彼此的优势，尤其需要考虑不同情境与文化对道德本质认知的调节作用(Lees & Gino, 2017)，并尽可能采用生态效度高的现实情境(Schein, 2020)，从而建构出更为完善的整合式理论模型。正如一元论者在最新研究中已经将情感因素纳入到以伤害为中心的道德判断模型中，试图化解对应道德理论与社会直觉模型的对立局面(Gray et al., in press)。同样地，多元论者也应在后续研究中继续厘清不同道德基础的本质区别，尽可能依据“奥卡姆剃刀原则”对原有理论删繁就简，并探究是否存在与以往研究不同质的道德基础，从而建构更为简洁和精确的理论模型。

最后需要说明的是，虽然道德多元论强调文化敏感性，道德一元论强调文化普适性，但是绝大多数实证研究却是以“怪异(WEIRD，指西方、受教育、工业化、富裕及民主)”样本为主导(Hester & Gray, 2020; Matsuo & Brown, in press)。而面对“非怪异”的样本，道德基础理论与对应道德理论往往缺乏解释力，如存在前人未曾发现的道德维度(e.g., Atari et al., 2020; Willard & McNamara, 2019)，以及不同的政治倾向与道德基础的对应模式(e.g., Davis et al., 2016; Frimer, 2020)，相对应的道德测量工具也存在构念效度低以及跨文化信效度不合格等问题(e.g., Iurino & Saucier, 2020; Zakharin & Bates, 2021)。未来研究应亟需考虑在不同文化下开展道德研究的可能性，尤其是针对中国本土文化的道德本质探究。中国哲学有诸多对道德本质的探讨，且相较于西方之求真，中国人更以求善的道德为终极追求（金观涛, 2021; 喻丰, 2021）。孔子提倡“三纲五常”，实则视一元之“仁”为终极道德境界（劳思光, 2019）。孟子所谓的“四端”看似道德多元论，然皆来源于“不忍人之心”此一元之道德心理情感基础（李泽厚, 2021）。同样地，宋代理学家朱熹主张的“存天理”与明代大儒王阳明提出的“致良知”亦是从“四端”中抽象概括出的一元论道德观（黎靖德, 2020）。由此可见，中国哲学从孔子到王阳明均体现出道德一元论特征。然而，中国人尤其崇尚整体思维与辩证思维，认为凡是相对相待的东西都不是绝然的对立，而是可以互相转化的（彭华，2017）。也就是说，一元论与多元论之争并不意味着两者是截然对立的：绝对的一元道德*本质*（比如“仁爱”）在

不同情境与文化下可以体现为不同的具体道德*现象*，如公平、关怀等，但无论是何种情况，都是为了遵循“仁爱”这种道德价值。所以以中国人的传统思维方式来看，可以同时容忍道德层面的“一”与“多”，两者可以相互转化与共存。最后，从本土心理学的视角来看，中国的“关系主义”文化又与西方的个体主义文化不同，导致中国人的道德以人际责任为基础，而不是个人权利（黄光国, 2006；杨中芳, 2009），这些因素均可能导致中国人对道德本质的认知与西方有实质差异。正如有研究发现中国人会认为违背“文明(civility)”的行为是不道德的(Buchtel et al., 2015)，而不是西方人认为的伤害或其他因素，并且“不文明”与“不道德”之间并无明显的内涵区分(Berniūnas et al., 2021)。以上思考也许均能为解决道德一元论与多元论的问题提供中国方案。

参考文献

黄光国. (2006). *儒家关系主义：文化反思与典范重建* (pp. 109–139). 北京：北京大学出版社.

金观涛. (2021). *轴心文明与现代社会*. 北京：东方出版社.

劳思光. (2019). *新编中国哲学史*. 北京：生活.读书.新知三联书店.

黎靖德. (主编). (2020). *朱子语类*(卷八十二). 北京：中华书局.

李泽厚. (2021). *中国古代思想史论*. 北京：人民文学出版社.

彭华. (2017). 中国传统思维的三个特征: 整体思维, 辩证思维, 直觉思维. *社会科学研究,* (3), 126–133.

亚里士多德. (2003). *尼各马可伦理学* (廖申白 译). 北京: 商务印书馆.

杨中芳. (2009). 中国人真是不用普遍性原则的吗：对道德思考发展研究的本土反思. 见 杨中芳 (编), *如何研究中国人* (pp. 107–137). 重庆：重庆大学出版社.

喻丰. (2021). 中西方思维究竟有何差异？. *山西师大学报(社会科学版), 48*(2), 20–26.

詹泽, 吴宝沛. (2019). 无处不在的伤害: 二元论视角下的道德判断. *心理科学进展, 27*(1), 128–140.

Alper, S., & Yilmaz, O. (2020). Does an abstract mind-set increase the internal consistency of moral attitudes and strengthen individualizing foundations? *Social Psychological and Personality Science*, *11*(3), 326–335.

Amin, A. B., Bednarczyk, R. A., Ray, C. E., Melchiori, K. J., Graham, J., Huntsinger, J. R., & Omer, S. B. (2017).

Association of moral values with vaccine hesitancy. *Nature Human Behaviour*, *1*(12), 873–880.

Atari, M., Graham, J., & Dehghani, M. (2020). Foundations of morality in Iran. *Evolution and Human Behavior*, *41*(5), 367–384.

Beal, B. (2020). What are the irreducible basic elements of morality? A critique of the debate over monism and pluralism in moral psychology. *Perspectives on Psychological Science*, *15*(2), 273–290.

Berniūnas, R., Silius, V., & Dranseika, V. (2021). Moralization East and West: Moralizing different transgressions in Chinese, Americans, and Lithuanians. *Asian Journal of Social Psychology*. Advance Online Publication.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., & Bond, M. H. (2015). Immorality East and West: Are immoral behaviors especially harmful, or especially uncivilized? *Personality and Social Psychology Bulletin*, *41*(10), 1382–1394.

Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review*, *19*(4), 371–394.

Chakroff, A., Russell, P. S., Piazza, J., & Young, L. (2017). From impure to harmful: Asymmetric expectations about immoral agents. *Journal of Experimental Social Psychology*, *69*, 201–209.

Curry, O. S., Chesters, M. J., & Van Lissa, C. J. (2019). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *Journal of Research in Personality*, *78*, 106–124.

Copes, H., Hochstetler, A., & Forsyth, C. J. (2013). Peaceful warriors: Codes for violence in adult male bar fighters. *Criminology, 51*, 761–794.

Davis, D. E., Rice, K., Van Tongeren, D. R., Hook, J. N., DeBlaere, C., Worthington, E. L., & Choe, E. (2016). The moral foundations hypothesis does not replicate well in Black samples. *Journal of Personality and Social Psychology*, *110*(4), e23–e30.

De Villiers-Botha, T. (2020). Haidt et al.'s case for moral pluralism revisited. *Philosophical Psychology*, *33*(2), 244–261.

Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., Vaisey, S., Iliev, R., & Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, *145*(3), 366–375.

DiBianca Fasoli, A. (2018). From autonomy to divinity: The cultural socialization of moral reasoning in an evangelical Christian community. *Child Development*, *89*(5), 1657–1673.

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences*, *151*, 109547.

Dungan, J. A., Chakroff, A., & Young, L. (2017). The relevance of moral norms in distinct relational contexts: Purity versus harm norms regulate self-directed actions. *PLoS One*, *12*(3), e0173405.

Durkheim, É. (1961). *Moral education*. (E. K. Wilson & H. Schnurer, Trans.). Cambridge: The Free Press. (Original work published 1925).

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*, 689–723.

Fiske, A. P. (1993). Social errors in four cultures: Evidence about universal forms of social relations. *Journal of Cross-Cultural Psychology, 24*(4), 463–494.

Franchin, L., Geipel, J., Hadjichristidis, C., & Surian, L. (2019). Many moral buttons or just one? Evidence from emotional facial expressions. *Cognition and Emotion*, *33*(5), 943–958.

Frimer, J. A. (2020). Do liberals and conservatives use different moral languages? Two replications and six extensions of Graham, Haidt, and Nosek's (2009) moral text analysis. *Journal of Research in Personality*, *84*, 103906.

Fodor, J. A. (1983). *The modularity of mind* (pp. 47–101). MIT press.

Giessner, S., & Van Quaquebeke, N. (2010). Using a relational models perspective to understand normatively appropriate conduct in ethical leadership. *Journal of Business Ethics, 95*, 43–55.

Gilligan, C. (1988). Remapping the moral domain: New images of self in relationship. In C. Gilligan, J. V. Ward & J. M. Taylor (Eds), *Mapping the Moral Domain* (pp. 3–19). Harvard University Press.

Graham, J. (2015). Explaining away differences in moral judgment: Comment on Gray and Keeney (2015). *Social Psychological and Personality Science*, *6*(8), 869–873.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology* (Vol. 47, pp. 55–130). Academic Press.

Graham, J., Haidt, J., Motyl, M., Meindl, P., Iskiwitch, C., & Mooijman, M. (2018). Moral foundations theory: On the advantages of moral pluralism over moral monism. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 211–222). Guilford Publications.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046.

Graham, J., & Iyer, R. (2012). The unbearable vagueness of "essence": Forty-four clarification questions for Gray, Young, and Waytz. *Psychological Inquiry*, *23*(2), 162–165.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619–619.

Gray, K. (2017). How to map theory: Reliable methods are fruitless without rigorous theory. *Perspectives on Psychological Science*, *12*(5), 731–741.

Gray, K., DiMaggio, N., Schein, C., & Kachanoff, F. (2021). What is "purity"? Conceptual murkiness in moral

psychology. *PsyArXiv.*

Gray, K., & Keeney, J. E. (2015a). Disconfirming moral foundations theory on its own terms: Reply to Graham (2015). *Social Psychological and Personality Science*, *6*(8), 874–877.

Gray, K., & Keeney, J. E. (2015b). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*, *6*(8), 859–868.

Gray, K., MacCormack, J., Henry, T., Banks, E., Schein, C., Armstrong-Carter, E., Abrams, S & Muscatell, K. (in press). The affective harm account (AHA) of moral judgment: Reconciling cognition and affect, dyadic morality and disgust, harm and purity. *Journal of Personality and Social Psychology*.

Gray, K., & Schein, C. (2016). No absolutism here: Harm predicts moral judgment 30× better than disgust—Commentary on Scott, Inbar, & Rozin (2016). *Perspectives on Psychological Science*, *11*(3), 325–329.

Gray, K., Schein, C., & Cameron, C. D. (2017). How to think about emotion and morality: Circles, not arrows. *Current Opinion in Psychology*, *17*, 41–46.

Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143*(4), 1600–1615.

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, *23*(2), 206–215.

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*(2), 101–124.

Guglielmo, S. (2018). Unfounded dumbfounding: How harm and purity undermine evidence for moral dumbfounding. *Cognition*, *170*, 334–337.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion* (pp. 123). New York: Pantheon Books.

Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Unpublished Manuscript, University of Virginia*, 191–221.

Haidt, J., Graham, J., & Ditto, P. (2015, October 27). *A straw man can never beat a shapeshifter: Response to Schein and Gray (2015).* Retrieved from http://www.yourmorals.org/blog/2015/10/a-straw-man-can-never-beat-a-sh apeshifter/

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*(4), 613–628.

Harper, C. A., & Rhodes, D. (2021). Reanalysing the factor structure of the moral foundations questionnaire. *British Journal of Social Psychology*, *60*, 1303–1329.

Hartman, R., Blakey, W., & Gray, K. (2022). Deconstructing moral character judgments. *Current Opinion in Psychology*, *43*, 205–212.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry*, *27*(1), 1–17.

Haslam, N., & Fiske, A. (1999). Relational models theory: A confirmatory factor analysis. *Personal Relationships, 6,* 241–250.

Hester, N., & Gray, K. (2020). The moral psychology of raceless, genderless strangers. *Perspectives on Psychological Science*, *15*(2), 216–230.

Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the Moral Foundations Questionnaire across 27

countries. *Assessment*, *27*(2), 365–372.

Kivikangas, J. M., Fernández-Castilla, B., Järvelä, S., Ravaja, N., & Lönnqvist, J.-E. (2021). Moral foundations and political orientation: Systematic review and meta-analysis. *Psychological Bulletin*, *147*(1), 55–94.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago: Rand McNally college publishing company.

Koleva, S., & Haidt, J. (2012). Let's use Einstein's safety razor, not Occam's Swiss army knife or Occam's chainsaw. *Psychological Inquiry*, *23*(2), 175–178.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, *46*(2), 184–194.

Landy, J. F. (2016). Representations of moral violations: Category members and associated features. *Judgment and Decision Making*, *11*(5), 496–508.

Lees, J., & Gino, F. (2017). Is the moral domain unique? A social influence perspective for the study of moral cognition. *Social and Personality Psychology Compass, 11*(8), e12327.

MacIntyre, A. (1957). What morality is not. *Philosophy*, *32*(123), 325–335.

Matsuo, A., & Brown, C. M. (in press). Culture points the moral compass: Shared basis of culture and morality. *Culture and Brain*.

McAdams, D. P., Albaugh, M., Farber, E., Daniels, J., Logan, R. L., & Olson, B. (2008). Family metaphors and moral intuitions: How conservatives and liberals narrate their lives. *Journal of Personality and Social Psychology*, *95*(4), 978–990.

McGraw, A. P., Schwartz, J. A., & Tetlock, P. E. (2012). From the commercial to the communal: Reframing taboo trade-offs in religious and pharmaceutical marketing. *Journal of Consumer Research, 39*, 157–173.

Menon, U., & Shweder, R. A. (1994). Kali's tongue: Cultural psychology and the power of shame in Orissa, India. In S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 241–282). American Psychological Association.

Piazza, J., Sousa, P., Rottman, J., & Syropoulos, S. (2019). Which appraisals are foundational to moral judgment? Harm, injustice, and beyond. *Social psychological and personality science, 10*(7), 903–913.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57–75.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard university press.

Rosenfeld, D. L., & Tomiyama, A. J. (2021). Moral judgments of COVID-19 social distancing violations: The roles of perceived harm and impurity. *Personality and Social Psychology Bulletin*, Advanced Online Publication.

Rottman, J., Kelemen, D., & Young, L. (2014). Tainting the soul: Purity concerns predict moral judgments of suicide. *Cognition*, *130*(2), 217–226.

Rottman, J., Kelemen, D., & Young, L. (2015). Hindering harm and preserving purity: How can moral psychology save the planet? *Philosophy Compass*, *10*(2), 134–144.

Rottman, J., & Young, L. (2019). Specks of dirt and tons of pain: Dosage distinguishes impurity from harm. *Psychological Science*, *30*(8), 1151–1160.

Royzman, E. B., & Borislow, S. H. (2022). The puzzle of wrongless harms: Some potential concerns for dyadic morality and related accounts. *Cognition, 220*, 104980.

Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, *10*(4), 296–313.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of*

Personality and Social Psychology, 76(4), 574–586.

Schein, C. (2020). The importance of context in moral judgments. *Perspectives on Psychological Science, 15*(2), 207–215.

Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, *41*(8), 1147–1163.

Schein, C., & Gray, K. (2016). Moralization and harmification: The dyadic loop explains how the innocuous becomes harmful and wrong. *Psychological Inquiry*, *27*(1), 62–65.

Schein, C., & Gray, K. (2018a). Moralization: How acts become wrong. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 363–370). Guilford Publications.

Schein, C., & Gray, K. (2018b). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70.

Schein, C., Hester, N., & Gray, K. (2016). The visual guide to morality: Vision as an integrative analogy for moral experience, variability and mechanism. *Social and Personality Psychology Compass*, *10*(4), 231–251.

Schein, C., Ritter, R. S., & Gray, K. (2016). Harm mediates the disgust-immorality link. *Emotion*, *16*(6), 862– 876.

Scott, S. E., Inbar, Y., & Rozin, P. (2016). Evidence for absolute moral opposition to genetically modified food in the United States. *Perspectives on Psychological Science*, *11*(3), 315–324.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The “big three” of morality (autonomy, community, divinity) and the “big three” explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119–169). New York, NY: Taylor & Frances/Routledge.

Stanley, M. L., Yin, S., & Sinnott-Armstrong, W. (2019). A reason-based explanation for moral dumbfounding. *Judgment and Decision Making*, *14*(2), 120–129.

Taylor, P. W. (1978). On taking the moral point of view. *Midwest Studies in Philosophy*, *3*(1), 35–61.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.

Uhlmann, E. L., & Zhu, L. (2014). Acts, persons, and intuitions: Person-centered cues and gut reactions to harmless transgressions. *Social Psychological and Personality Science*, *5*(3), 279–285.

Wagemans, F., Brandt, M. J., & Zeelenberg, M. (2018). Disgust sensitivity is primarily associated with purity-based moral judgments. *Emotion*, *18*(2), 277–289.

Wasserman, E. A., Chakroff, A., Saxe, R., & Young, L. (2017). Illuminating the conceptual structure of the space of moral violations with searchlight representational similarity analysis. *NeuroImage*, *159*, 371–387.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological science,* *16*(10), 780–784.

Willard, A. K., & McNamara, R. A. (2019). The minds of god (s) and humans: Differences in mind perception in Fiji and North America. *Cognitive Science*, *43*(1), e12703.

Wisneski, D. C., & Skitka, L. J. (2017). Moralization through moral shock: Exploring emotional antecedents to moral conviction. *Personality and Social Psychology Bulletin*, *43*(2), 139–150.

Yalçındağ, B., Özkan, T., Cesur, S., Yilmaz, O., Tepe, B., Piyale, Z. E., Biten, A. F., & Sunar, D. (2019). An investigation of moral foundations theory in Turkey using different measures. *Current Psychology*, *38*(2), 440–457.

Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, *120*(2), 202–214.

Zakharin, M., & Bates, T. (2021). Remapping the foundations of morality: Well-fitting structural model of the Moral Foundations Questionnaire. *PsyArXiv*..

Monism and pluralism in morality: Origins, connotations and debates

ZHANG Haotian[1], YU Feng[1], XU Liying[2], XUAN Zheli[1]

([1]*Department of Psychology, School of Philosophy, Wuhan University, Wuhan 430072, China*)

([2] *School of Social Marxism, Tsinghua University, Beijing 100084, China*)

Abstract: In recent years, the debates between monism and pluralism are one of the most active theoretical disagreements in the field of moral psychology. Moral monism claims that all moral phenomena on the surface and moral structures behind them can be explained by one factor, and the representative theories are stages theory of moral development and dyadic morality theory, etc. Moral pluralism holds that morality can't be explained by a single factor, whereas has many heterogeneous moral dimensions and cultural sensitivity. The representative theories are triadic moral discourse theory, relational model theory and moral foundations theory, etc. Moreover, moral foundations theory and dyadic morality theory are the typical representatives of the debates between monism and pluralism. The two theories have engaged in a long and inconclusive dialogue on the harm, purity, modular claims, and moral foundations of politics. Future studies should further explore the monism and pluralism of morality from three specific aspects so as to maintain the vitality of theory in the field of moral psychology.

Key words: Monism, Pluralism, Moral Foundations Theory, Dyadic Morality Theory, Haidt-Gray Debates